CASE NO. 2023-1663

NONCONFIDENTIAL VERSION

BEFORE THE UNITED STATES
COURT OF APPEALS FOR THE FEDERAL CIRCUIT

STUPP CORPORATION, A DIVISION OF STUPP BROS, INC.,
IPSCO TUBULARS, INC., MAVERICK TUBE CORPORATION,

*Plaintiffs,*

WELSPUN TUBULAR LLC USA,

*Plaintiff-Appellee*

V.

UNITED STATES,

*Defendant-Appellee,*

HYUNDAI STEEL COMPANY,

*Defendant,*

SEAH STEEL CORP.,

*Defendant-Appellant*

APPEAL FROM THE U.S. COURT OF INTERNATIONAL TRADE
IN CASE NOS. 15-CV-00334, 15-CV-00336, AND 15-CV-00337,
JUDGE CLAIRE R. KELLY

NONCONFIDENTIAL REPLY BRIEF
OF APPELLANT SEAH STEEL CORPORATION

*Submitted by*

Jeffrey M. Winton
Jooyoun Jeong

WINTON & CHAPMAN PLLC
1100 13th Street, N.W., Suite 825
Washington, D.C. 20005
(202) 774-5500

Attorneys for SeAH Steel Corporation

January 8, 2024

# UNITED STATES COURT OF APPEALS
# FOR THE FEDERAL CIRCUIT

## CERTIFICATE OF INTEREST

**Case Number**   2023-1663

**Short Case Caption**   Stupp Corporation v. U.S.

**Filing Party/Entity**   SeAH Steel Corporation

---

**Instructions:**

1. Complete each section of the form and select none or N/A if appropriate.

2. Please enter only one item per box; attach additional pages as needed, and check the box to indicate such pages are attached.

3. In answering Sections 2 and 3, be specific as to which represented entities the answers apply; lack of specificity may result in non-compliance.

4. Please do not duplicate entries within Section 5.

5. Counsel must file an amended Certificate of Interest within seven days after any information on this form changes.  Fed. Cir. R. 47.4(c).

---

I certify the following information and any attached sheets are accurate and complete to the best of my knowledge.

Date: 01/08/2024

Signature:   /s/ Jeffrey M. Winton

Name:   Jeffrey M. Winton

**FORM 9. Certificate of Interest**

| 1. Represented Entities. Fed. Cir. R. 47.4(a)(1). | 2. Real Party in Interest. Fed. Cir. R. 47.4(a)(2). | 3. Parent Corporations and Stockholders. Fed. Cir. R. 47.4(a)(3). |
|---|---|---|
| Provide the full names of all entities represented by undersigned counsel in this case. | Provide the full names of all real parties in interest for the entities. Do not list the real parties if they are the same as the entities. <br><br> ☑ None/Not Applicable | Provide the full names of all parent corporations for the entities and all publicly held companies that own 10% or more stock in the entities. <br><br> ☐ None/Not Applicable |
| SeAH Steel Corporation | | SeAH Steel Holdings Corporation |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |

☐  Additional pages attached

**4. Legal Representatives.** List all law firms, partners, and associates that (a) appeared for the entities in the originating court or agency or (b) are expected to appear in this court for the entities. Do not include those who have already entered an appearance in this court. Fed. Cir. R. 47.4(a)(4).

☐    None/Not Applicable          ☐    Additional pages attached

| | | |
|---|---|---|
| Winton & Chapman PLLC | Law Office of Jeffrey M. Winton PLLC | |
| Amrietha Nellan | | |
| | | |

**5. Related Cases.** Other than the originating case(s) for this case, are there related or prior cases that meet the criteria under Fed. Cir. R. 47.5(a)?

☑ Yes (file separate notice; see below)   ☐ No   ☐ N/A (amicus/movant)

If yes, concurrently file a separate Notice of Related Case Information that complies with Fed. Cir. R. 47.5(b). **Please do not duplicate information.** This separate Notice must only be filed with the first Certificate of Interest or, subsequently, if information changes during the pendency of the appeal. Fed. Cir. R. 47.5(b).

**6. Organizational Victims and Bankruptcy Cases**. Provide any information required under Fed. R. App. P. 26.1(b) (organizational victims in criminal cases) and 26.1(c) (bankruptcy case debtors and trustees). Fed. Cir. R. 47.4(a)(6).

☑    None/Not Applicable          ☐    Additional pages attached

| | | |
|---|---|---|
| | | |
| | | |

Table of Contents

CONFIDENTIAL MATERIAL OMMITTED

Respondent's confidential sales information is omitted from lines 2, 3, and 4 on page 26.

Diagrams representing Respondent's confidential sales information are omitted from pages 26 and 27.

Table of Authorities

Page

NONCONFIDENTIAL REPLY BRIEF OF APPELLANT
SEAH STEEL CORPORATION

This brief is submitted on behalf of Appellant SeAH Steel Corporation

("SeAH") in the appeal of the February 24, 2023, decision by the Court of

International Trade (the "CIT") in *Stupp v. U.S.,* Consol. Ct. No. 15-00334.

A. *Commerce's Use of Professor Cohen's Proposed*
*Thresholds for Evaluating the d Statistic Is Not*
*Consistent with Normal Statistical Practice*

In its decision in *Stupp III,* this Court held that, when an agency purports

to apply a mathematical tool, it must apply that tool in a manner consistent

with the tool's assumptions.[1]  As the Court explained,

> ... Professor Cohen derived his interpretive cutoffs under
> certain assumptions. Violating those assumptions can
> subvert the usefulness of the interpretive cutoffs,
> transforming what might be a conservative cutoff into a
> meaningless comparator.[2]

In *Stupp III,* the Court concluded that "the evidence and arguments before us

call into question whether Commerce's application of the Cohen's *d* test to

the data in this case violated the assumptions of normality, sufficient

observation size, and roughly equal variances associated with that test."[3]  This

---

[1] *Stupp Corp. v. U.S.*, 5 F.4th 1341 (Fed.Cir. 2021) (hereinafter "*Stupp III*").

[2] *See Stupp III* at 1360, citing *Virnetx, Inc. v. Cisco Sys., Inc.*, 767 F.3d 1308,
1332 (Fed.Cir. 2014).

[3] *Id.*

Court therefore remanded this case to give Commerce an opportunity to explain "whether the limits on the use of the Cohen's $d$ test prescribed by Professor Cohen and other authorities were satisfied in this case or whether those limits need not be observed when Commerce uses the Cohen's $d$ test in less-than-fair-value adjudications."[4]  And, the Court explicitly invited Commerce "to clarify its argument that having the entire universe of data rather than a sample makes it permissible to disregard the otherwise-applicable limitations on the use of the Cohen's $d$ test."[5]

As explained in our initial brief, Commerce's redetermination failed to address the Court's concerns.  Commerce did not explain how its use of Cohen's $d$ was consistent with normal statistical practice.  Instead, Commerce asserted that it was not bound by normal statistical practice, because it was not using Cohen's $d$ in the manner that statisticians use it.[6]  Commerce's assertion means that it cannot rely on statistical practice to justify its use of Cohen's $d$ in the DPA.  Instead, if the DPA is to be upheld, Commerce must demonstrate that its use of Cohen's $d$ in a manner inconsistent with statistical practice is nevertheless reasonable.

---

[4] *Id.*

[5] *Id.*

[6] *See* SeAH Initial Brief at 8-9.

In this regard, we note that the Government's brief repeatedly cites to this

Court's 2019 decision in *Mid Continent Steel & Wire* as upholding

Commerce's use of Cohen's *d* against SeAH's challenges.[7]  But that *Mid*

*Continent* decision pre-dated the decision in *Stupp III* that led to the remand

that is the subject of the present proceedings.  This Court undoubtedly was

aware of the *Mid Continent* when it issued *Stupp III*.  If *Mid Continent* had

disposed of SeAH's arguments, we would not be here.

> 1. *Nothing in the Academic Literature Addressed by Commerce Authorizes the Use of Cohen's d when the Underlying Assumptions Are Not Satisfied, Whether the Analysis Concerns an Entire Population or Just a Sample*

The Government asserts that Commerce's methodology can be

considered consistent with normal statistical practice, because Professor

Cohen's text states that "effect size is a phenomenon of the population and it

exists independent of any statistical analysis based on data sampled from that

population."[8]  Based on that statement, the Government contends that

Commerce was permitted to disregard the assumptions required by Professor

---

[7] *See* Government Brief at 41, 44, and 47, citing *Mid Continent Steel & Wire, Inc. v. U.S.*, 940 F.3d 662 (Fed.Cir. 2019).

[8] *Id.* at 26.

- 3 -

Cohen's text for use of the *d* statistic, because its analysis considered the full population, and not just a sample.[9]

The Government's argument is, however, fundamentally dishonest. To begin with, Professor Cohen's reference to effect size as "the degree to which the phenomenon is present in the population" is found in Chapter 1 of his text, where Professor Cohen introduces the general concept of "power analysis." It is not found in Chapter 2, which describes the use of the *d* statistic. Professor Cohen explicitly stated that "Each of the Chapters 2-10 will present in some detail the ES index appropriate to the test to which the chapter is devoted."[10] Chapter 2 is devoted to the t-test. And, as Professor Cohen made clear,[11] and as Commerce has admitted,[12] a t-test is appropriate only with data that is Normally distributed and has equal variances and sufficient data points. Professor Cohen's general statements about effect sizes in a population in Chapter 1 of his text simply have no bearing on the proper use of the *d* statistic, as described in Chapter 2 of his book.

---

[9] *Id.*

[10] *See* Cohen, STATISTICAL POWER ANALYSIS FOR THE BEHAVIORAL SCIENCES (2d ed. 1988) at 13 (Appx3757).

[11] *Id.,* at 19 (Appx3762).

[12] *See* Commerce's April 4 Final Results of Redetermination Pursuant to Court Remand at 42-43 (Appx0069-0070) (hereinafter "Redetermination").

More generally, the description of the "effect size" as the "degree to which the phenomenon is present in the population" says nothing about the use of any particular test or thresholds to evaluate the effect size in the population.  The fact that a population may have an inherent "effect size" does not mean that it is appropriate to rely on a $d$ statistic, and the specific thresholds that Professor Cohen proposed for use with that statistic, when Professor Cohen explicit stated that the $d$ statistic was intended *only* for use in conjunction with a t-test on data that is Normally distributed and has equal variances and sufficient data points.

As we have explained previously, Normal distributions — whether based on a sample or encompassing an entire population — have very different mathematical characteristics than non-Normal distributions.[13]  In light of the mathematical differences between Normal and non-Normal distributions, there is no reason to believe that a rule-of-thumb (such as Professor Cohen's proposed thresholds) that was developed to be used with Normal distributions can properly be applied to non-Normal data — whether the data being considered is a sample or an entire population.  Commerce has never offered any mathematical calculations supporting the conclusion that Professor

---

[13] *See* SeAH Initial Brief at 26-27.

Cohen's proposed thresholds have the same meaning for non-Normal data

that they would for Normal distributions.  Nothing in the briefs submitted by

the Government or Welspun remedy that fundamental omission.

> 2.  *The Government's Suggestion that SeAH*
>     *Made a "Concession" that the Academic*
>     *Literature Does Not Preclude Commerce's*
>     *Use of Cohen's d Is False and Misleading*

In our initial brief, we noted that Commerce's argument regarding the use

of Cohen's *d* was "purely negative," because Commerce asserted only that

none of the texts on the record prohibit its proposed use of Cohen's *d*.  In

response, the Government asserts that our statement represents "a concession"

that "nothing in the academic literature states that it is *improper* to use

Cohen's *d* in the manner utilized by Commerce."[14]

As discussed above, the Government's argument is directly contradicted

by Professor Cohen's own text, which specifically states that the *d* statistic

was intended only to be used in conjunction with a t-test, and that t-tests may

be used only when the data satisfies the assumptions of Normal distribution,

equal variances, and sufficient data points.[15]  More generally, the

---

[14] Government Brief at 23.

[15] *See e.g., Cohen* at 17, 19 (Appx 3761-3762).

Government's assertion ignores the very limited nature of the remand proceeding conducted by Commerce.

In Commerce's remand proceeding, SeAH *was not permitted* to introduce any and all academic materials it deemed relevant.  Instead, Commerce allowed SeAH only to place on the record the academic materials that had previously been cited to this Court.  Commerce then attempted to explain why none of the material previously cited to this Court addressed its use of Cohen's *d*.  But Commerce never essayed a full survey of the academic literature, and it did not permit SeAH to present such a survey or other evidence of the academic understanding of the proper use of Cohen's *d*.

As the Canadian *amici* have noted, there are other sources of information about the academic understanding of the limitations on the use of Cohen's *d* that might have been made available to Commerce if Commerce had permitted a more wide-ranging proceeding.  For example, the Canadian *amici* have, in other proceedings, submitted an export report by Professor Larry Hedges addressing the consistency of Commerce's use of Cohen's *d* with the academic understanding.[16]

_____

[16] Amici Brief at 3 (citing Larry V. Hedges, *Review and Analysis of the Cohen's* d *Test as Used in the U.S. Department of Commerce's Differential Pricing Methodology* (Dec. 27, 2022) (hereinafter "Hedges' Report")).

Professor Hedges is a world-renowned expert on "effect-size" analysis, whose *g* statistic for measuring effect size "was ... developed to remove a small positive bias affecting the calculation of *d*,"[17] and is "commonly used."[18]  His conclusions regarding Commerce's use of Cohen's *d* are clear:

> The relations between Cohen's *d* and measures of overlap used to interpret *d* do not hold when the distributions being compared do not have equal standard deviations or are not normally distributed. Therefore, any interpretations of *d* based on assuming normal distributions with equal standard deviations cannot be relied upon if these assumptions are violated....[19]

> When the assumption that the two populations being compared are normally distributed and have the same standard deviation are not met, Cohen's *d* cannot be the interpreted in the same way as when these assumptions are met. Indeed, *Cohen never suggested using d to compare distributions that were not normal and it would not be accepted statistical practice to do so*.[20]

> Regardless of whether the distributions being compared are samples or populations, the interpretation of *d* remains the same—when the distributions are normal with equal

---

[17] *See* Ellis, Paul, THE ESSENTIAL GUIDE TO EFFECT SIZES: STATISTICAL POWER, META-ANALYSIS, AND THE INTERPRETATION OF RESEARCH RESULTS (2010) at 27 (Appx4411).

[18] *See* Lane, David, *et al.*, *Introduction to Statistics, Online Edition*, Chapter 19 (Appx4350-4351).

[19] Hedges' Report at 3-4.

[20] *Id.* at 15 (emphasis added).

standard deviations the interpretation of *d* is different than when they are not….[21]

Whether the groups Commerce uses to compute *d* are a population or a sample, the interpretation of *d* given by Cohen and other statistical experts can only be relied upon if the assumptions of normality and equal standard deviations are satisfied.[22]

Commerce asserts that its analyses are all population based. Even if this assertion were correct, it would not obviate the need to establish that their effect sizes are large enough to be of practical importance. *Because Cohen's conventions assume that the distributions being compared are normal with the same standard deviation, those conventions cannot be used to establish practical importance of d values computed from population or sample data in which these assumptions are violated.*[23]

As Professor Hedges amply demonstrated in his report, Commerce's use of Cohen's *d* is not consistent with common statistical practice.

Of course, we cannot fault Commerce for failing to address an academic review of its methodology that was not included in the record before it. But Professor Hedges' report clearly demonstrates the limitations of Commerce's blinkered remand proceeding, which failed to permit a broad discussion of the academic understanding of Cohen's *d*. And it puts the lie to the

---

[21] *Id.,* Appendix II, at v.

[22] *Id.,* Appendix II, at ix.

[23] *Id.,* Appendix II, at xviii (emphasis added).

Government's assertion that our inability to cite such materials (due to

Commerce's restrictions on the remand proceeding) constitutes a

"concession" that no academic literature disagrees with Commerce's use of

Cohen's *d.*

We understand that the Government has objected to citations to non-

record academic materials in the initial briefs by SeAH and the Canadian

*amici.*[24]  The academic materials on which *amici* rely are not adjudicative

facts that constitute the "facts of the particular case" such that they must

either be on the record or subject to judicial notice. They instead reflect

principles of statistics bearing on Commerce's generally applicable

methodology.[25]  Indeed, this Court has emphasized the importance of the

statistics literature in evaluating Commerce's use of Cohen's d and has

considered statistics literature not found on the administrative record in

*Stupp III.*[26]  There does not appear to be any reason for the Court to take a

different approach here.

---

[24] Government Brief at 60-61.

[25] *See* Fed. R. Evid. 201, cmt. (a).

[26] *See Mid Continent Steel & Wire, Inc. v. U.S.*, 31 F.4th 1367, 1381 (Fed.Cir. 2022); *Stupp III*, 5 F.4th 1357-58.

Finally, we note that we have cited Professor Hedges' report only to demonstrate the falsity of the Government's suggestion that limitations imposed by Commerce on its remand somehow constitute a concession by SeAH that there is no academic material refuting Commerce's proposed use of Cohen's *d*.

### B. Commerce's Use of the Cohen's d Test Can Classify Imperceptible Differences as Significant, Just as this Court Predicted

In *Stupp III*, this Court expressed a separate concern that Commerce's simplistic application of its Cohen's *d* test could generate arbitrary results when applied to data with a small number of observations or small price differences.  To support this point, the Court provided a hypothetical example showing that, when the variances in the data are small, even tiny differences that had no practical significance could result in a "large" value for *d*.[27]

Commerce's Redetermination did not appear to dispute the conclusion that the example set forth in the *Stupp III* decision would generate am in correct "passing" result under the Cohen's *d* test.  However, the Redetermination asserted that any harm caused by that incorrect result would be ameliorated by the "meaningful difference" test, since the minor price

---

[27] *Stupp III* at 1358-59.

differences generated by the *Stupp III* decision's hypothetical example would result in an insignificant difference in the dumping margins under the different comparison methodologies.[28]

As we demonstrated in our initial brief, Commerce's explanation holds true only if there is only one product under consideration. If there are multiple products under consideration, an incorrect finding that there are "large" price differences for any one product will affect the ultimate outcome of Commerce's "Ratio Test," which depends on the percentage of sales with "large" price differences. Consequently, even if an incorrect finding of a "large" price difference for one product does not affect the dumping margin calculated for *that* product, it may affect the results of the "Ratio Test" and, hence, the dumping margins calculated for other products.

Our initial brief included a hypothetical example demonstrating how, in a two-product model, the "meaningful difference" test would not prevent dumping margins from being created by insignificant price differences for one product that "pass" the Cohen's *d* test.[29] In response, the Government contends that it can construct an alternative hypothetical example of a two-

---

[28] Redetermination at 31 (Appx0058).

[29] Opening Brief at 28-30.

product model in which dumping margins would not be created.[30]  But the

Government's hypothetical is irrelevant.  Our brief demonstrated that

Commerce was wrong when it asserted that its "meaningful difference" test

would *necessarily* prevent the creation of dumping margins from insignificant

price differences.

Of course, we can construct an infinite number of alternative hypothetical

examples in which dumping margins are or are not created.  The existence of

*any* examples in which dumping margins are created fully disproves

Commerce's claim that, because of the "meaningful difference" test, there is

no reason for concern about potentially distorted results.

C.  *Commerce's "Ratio Test" Cannot Give
        Meaningful Results when the Cohen's d Test Fails
        to Properly Identify Significant Price Differences*

1.  *If Sales Are Incorrectly Identified as "Passing"
        the Cohen's d Test, then Commerce's Calculation
        of the "Ratio Test" Will Necessarily Be Distorted*

By its terms, the results of Commerce's "Ratio Test" will depend on the

number of sales that "pass" the Cohen's $d$ Test.  If the results of the Cohen's $d$

test are inaccurate, the results of the "Ratio Test" will also be inaccurate.

---

[30] Government Brief at 31-33.

That conclusion follows, necessarily, from the manner in which the Ratio Test is calculated.

Both the Government and Welspun agree that the use of Cohen's *d* with data that does not satisfy Professor Cohen's assumptions may result in either an understatement or overstatement of the number of sales passing the Cohen's *d* test. It follows, then that any ratio calculated based on the number of sales that have a calculated *d* greater than 0.8 may also be distorted.

The Government asserts that any distortion caused by the improper results of the Cohen's *d* test is addressed by the fact that the Ratio Test requires a minimum finding that 33 percent of the sales "pass" the Cohen's *d* test before an alternative margin-calculation methodology is considered.[31] But, neither Commerce nor the government nor any of the academic literature quantify the precise effect of applying the Cohen's *d* test to data that does not satisfy Professor Cohen's assumptions. To the contrary, the academic literature suggests that the values of Cohen's *d*, when these assumptions are not met, are meaningless.[32]

---

[31] Government Brief at 35.

[32] *Stupp III*, 5 F.4th at 1360.

Furthermore, even if it were somehow possible to identify a "correct" pass rate in such circumstances,[33] a small distortion in the Cohen's *d* for any comparison may have a large impact on the pass rate. For example, if one somehow could determine that the "correct" Cohen's *d* would be 0.799 for every sale, an improper increase in Cohen's *d* of just 0.01 would result in all sales moving from not passing to passing — which would change the Ratio Test result from 0 percent to 100 percent. By the same token, if the "correct" result for the Ratio Test would be 32.999 percent, an improper increase of just 0.001 percent in the number of sales "passing" the Cohen's *d* test would move the Ratio Test result from a negative result to a positive result.

2. *If the Assumptions Required for Use of Cohen's d Are Not Satisfied, then the Ratio Test Cannot Properly Distinguish Actual "Patterns" from Random Chance*

As we have explained previously, Commerce has made clear that, in order to make a finding that a "pattern" exists that would justify a departure from the normal A-to-A comparison under the statute, it must distinguish between differences that have meaning and those that arise by chance.[34]

---

[33] We note that, when the data does not satisfy Professor Cohen's assumptions, it is not possible to identify a "correct" Cohen's *d* or a "correct" "pass rate" to be used in the Ratio Test.

[34] *See, e.g.,* SeAH's Initial Brief at 36.

When the assumptions described by Professor Cohen are satisfied, a "power analysis" using the *d* statistic allows a calculation of the likelihood that an observed difference reflects a true pattern, and not just chance fluctuations. By contrast, when those assumptions are not satisfied, no such calculation is possible.

The Government contends that the Cohen's *d* test is irrelevant to Commerce's detection of "patterns," because the Cohen's *d* test was intended *only* to determine whether an observed price difference is "significant," while the separate Ratio Test is the exclusive tool for determining whether a "pattern" exists.[35]  The Government further asserts that the use of the Ratio Test to distinguish between patterns and chance fluctuations was upheld by this Court in *Stupp III.*[36]

That argument suffers from two logical flaws:  First, as discussed above, the results of the Ratio Test are not independent from the results of the Cohen's *d* test.  Instead, because the "Ratio Test" simply counts the number of sales that "pass" the Cohen's *d* test, a flaw in the Cohen's *d* test necessarily results in an incorrect result from the Ratio Test.  If the Cohen's *d* test is

---

[35] *See* Government Brief at 25, 33-35.

[36] *See id.,* at 34.

flawed, then any finding regarding the existence of a pattern based on the "Ratio Test" will also be flawed.

Second, as we have explained previously, this Court's decision in *Stupp III* did not hold that the "Ratio Test" provided a reasonable tool for distinguishing between true patterns and chance fluctuations. Such an argument was never presented to the Court in *Stupp III*. Instead, SeAH's challenge to the Ratio Test in *Stupp III* was based on the claim that the Ratio Test used completely arbitrary cut-offs that had no basis in evidence, mathematics, or logic. The Court upheld Commerce's choice of cut-offs solely on the grounds that they were as reasonable as any other arbitrary cut-offs that it considered.[37]

In these circumstances, Commerce's use of Cohen's $d$ when the assumptions described by Professor Cohen are not satisfied means that its DPA fails to satisfy the statute's requirement that there be a "pattern" of significant price differences, and not just chance fluctuations, before Commerce may consider an alternative dumping-margin calculation methodology. Nothing in Commerce's redetermination and nothing in the

---

[37] *See Stupp III,* 5 F.4th at 1355. *See also* SeAH's Initial Brief at 35.

Appellees' briefs explains how "actual patterns" might be distinguished from random fluctuations using Commerce's DPA.

### 3. Record Evidence Confirms that the Results of the DPA in this Case Were Inflated by Random Changes in Exchange Rates that Were Outside SeAH's Control

The Government also contends that the failure of the DPA to distinguish "actual patterns" from random fluctuations is irrelevant, because "{t}he allegation that a company's prices may be set by random chance has no basis in law or fact."[38]  Tellingly, the Government does not cite any principle of *law* that prevents random factors from affecting a company's prices.  Nor does it provide any factual or logical justification for its claim:  Economic theory states that prices are set by the interplay of supply and demand in the market, not by the dictates of any one producer.  And the evidence is clear that SeAH does not have the kind of market power needed to dictate prices to its customers. [39]

---

[38] Government Brief at 34.

[39] The petition in this case indicates that SeAH sold its product during the relevant period in competition with 18 U.S. producers, and also faced competition from 12 other Korean producers and 6 Turkish producers, in addition to producers in other countries.  *See* Petition, Volume I, at 2-5 (listing U.S. producers) and Exhibit I-4 (listing Korean and Turkish producers) (Appx0551-0554, Appx0594-0600).  Furthermore, the evidence shows that the line pipe offered by these producers was fungible and sold in the same geographic markets.  *See id.*, at 20-21 (Appx0569-0570).

Furthermore, the evidence in this particular case demonstrates precisely how Commerce's results can be distorted by random factors. It is well-settled, for example, that exchange rate movements are indistinguishable from a "random walk" — and plainly not under the control of any particular exporter.[40] And, in this case, the slight movements in exchange rates over the investigation period were sufficient, by themselves, to create an affirmative finding under Commerce's DPA.

In our arguments before the CIT, we presented the following table which summarized the average exchange rate and the standard deviation for each individual quarter of the investigation period, the average exchange rate and standard deviation for the remaining three quarters of the period to compare to each quarter, the pooled standard deviation, and the Cohen's *d* statistic for comparisons of each quarter to the other three.[41]

---

[40] *See, e.g.,* F. Alvarez, A. Atkeson, and P. Kehoe, *If Exchange Rates Are Random Walks, Then Almost Everything We Say about Monetary Policy is Wrong*, FEDERAL RESERVE BANK OF MINNEAPOLIS RESEARCH DEPARTMENT STAFF REPORT 388, 1 (Mar. 2007) (reporting "well-established feature of the data" that "nominal rates of exchange between major currencies are well-approximated by random walks.").

[41] SeAH's June 14, 2022, Comments on Third Remand Redetermination at 24 (Appx5627).

Analysis of Exchange Rates by Quarter

| | | Q1 (10~12/2013) | Q2 (1~3/2014) | Q3 (4~6/2014) | Q4 (7~9/2014) |
|---|---|---|---|---|---|
| This Quarter | Minimum | 0.000910 | 0.000922 | 0.000937 | 0.000958 |
| | Maximum | 0.000952 | 0.000952 | 0.000987 | 0.000991 |
| | Average | 0.000940 | 0.000936 | 0.000966 | 0.000974 |
| | Standard Deviation | 0.000009 | 0.000006 | 0.000016 | 0.000009 |
| Other 3 Quarters | Average | 0.000959 | 0.000960 | 0.000950 | 0.000947 |
| | Standard Deviation | 0.000020 | 0.000019 | 0.000019 | 0.000018 |
| Comparison | Difference in Means | 0.000019 | 0.000025 | 0.000016 | 0.000027 |
| | Pooled Std. Dev. | 0.000016 | 0.000014 | 0.000018 | 0.000014 |
| | **Cohen's *d*** | **1.18915** | **1.72082** | **0.88579** | **1.91488** |

As that analysis demonstrated, the Cohen's *d* statistics for comparisons of the exchange rates in each quarter to the exchange rates in the other three quarters are greater than 0.8 for every single quarter.

Furthermore, we also demonstrated that the same result would occur *even if the exporter fixed its U.S. invoice prices at a uniform amount* in U.S. dollars, as long as there was some expense incurred in Korean Won that was subtracted from the invoice amount to determine the net U.S. price used in Commerce's analysis. For example, SeAH reported in this case that it incurred costs in Korean Won to transport merchandise from its plant to the Korean port of exportation. For sales made from U.S. inventory, SeAH could not identify the actual cost of this Korean inland freight on a transaction-specific basis, but instead reported a period-average cost of roughly 13,000

Korean Won per ton.[42]  Commerce converted that amount into U.S. dollars

for each sale using the exchange rate on the date of the sale, and subtracted

that U.S. dollar amount from SeAH's U.S. dollar invoice prices to determine

the net U.S. price used in its Cohen's *d* calculations.  As we demonstrated to

the CIT, the minor fluctuations in the exchange rates used to convert this

relatively minor expense in Korean Won to U.S. dollars would lead

Commerce to find that 100 percent of SeAH's sales "passed" the Cohen's *d*

test and thus apply the A-to-T methodology to all sales, even though the same

expense was reported for all U.S. sales, and *even if SeAH charged the same*

*price to all U.S. customers on all sales during the period*.[43]

The evidence therefore confirms that the analysis of price differences

over time was necessarily distorted by slight changes in exchange rates over

the investigation period.  Those exchange-rate changes were, by themselves,

sufficient to cause *all* of SeAH's U.S. sales to pass the Ratio Test.  Contrary

to the Government's claims, the evidence demonstrates that SeAH's prices

were affected by random factors, and the distortion in this case was not

simply hypothetical.

---

[42] *See* SeAH's February 2, 2015, Section C Response at 26-27 and
Appendix C-7 (Appx6341-6342, Appx6367-6369).

[43] *See* SeAH's Comments on Third Remand at 25-28 (Appx5628-5631).

D.  *Commerce's Assertion that Professor Cohen's Proposed Thresholds Can Be Used as Universal Yardsticks Because They Are Based on Real-World Observations Is Illogical and Contrary to the Evidence*

As noted in our initial brief, Commerce's Redetermination did not attempt to justify its use of Professor Cohen's proposed thresholds as a matter of statistical practice or as a matter of mathematical logic. Instead, Commerce fell back on the claim that Professor Cohen's rule-of-thumb that a *d* of 0.8 indicates a "large" effect can be used as a universal yardstick, because it reflects "real-world observed differences."[44] Indeed, Commerce's Redetermination concluded with the observation that "Dr. Cohen's thresholds are not based on the alleged statistical criteria but, rather, on real-world observations."[45]

It is, undoubtedly, true that Commerce is not required to follow statistical principles in its dumping calculations. But that does not mean that it is reasonable for Commerce to rely on any rule-of-thumb that is purportedly derived from real-world observations of some unrelated phenomenon. Instead, before Commerce can rely on such a rule-of-thumb, it must demonstrate that the phenomenon from which the rule-of-thumb was derived

---

[44] *See* Redetermination at 20, 24, 28 (Appx0047, 0051, 0055).

[45] *Id.* at 65 (Appx0092).

has some rational relationship to the situation for which Commerce intends to apply that rule. In this case, Commerce has absolutely failed to meet that burden.

### 1. Professor Cohen's Rules-of-Thumb Do Not Establish Universal Yardsticks

As explained in our initial brief, Professor Cohen made it absolutely clear that his proposed thresholds for evaluating $d$ were not intended to establish universal yardsticks that might be applied when the data being analyzed did not meet the assumptions of Normality, equal variances, and sufficient data points. Indeed, Professor Cohen cautioned against the use of his thresholds, even when his assumptions were satisfied, outside of the specific field of behavioral sciences.[46]

The Government asserts that Professor Cohen "expresses no such limitation."[47] In support of that claim, the Government points to Professor Cohen's statements that he hoped that his proposed cut-offs would be "easy to grasp" and "will be found reasonable by reasonable people."[48] But neither of those statements contradicts Professor Cohen's explicit statements that the $d$

---

[46] *See Cohen* at 25 (Appx3768).

[47] Government Brief at 42.

[48] *Id.*

statistic was to be used *only* in conjunction with a t-test and that a t-test can only be used when the assumptions of Normality, equal variances, and sufficient data points are satisfied. Furthermore, neither of the statements cited by the Government contradicts Professor Cohen's explicit statement that the evaluation of any particular value of *d* had to be "relative … to the area of behavioral science or even more particularly to the specific content and research method being employed in any given investigation."[49]

Finally, regardless of Professor Cohen's intent, it is simply unreasonable to assume that rules-of-thumb based on the heights of teenaged girls or the IQs of different types of students establish a yardstick for evaluating SeAH's prices in the absence of some reason to believe that prices for steel pipe should follow the same patterns as heights and IQs. As we have noted previously, it might be acceptable to use thresholds derived from an analysis of Normal distributions as a yardstick for evaluating other Normal distributions. And it is true that heights and IQs are Normally distributed. But there is no reason to believe that the prices for steel pipe should be

---

[49] *Cohen* at 25 (Appx3768).

Normally distributed, and the evidence on the record confirms that SeAH's

actual prices are not Normally distributed.[50]

Commerce certainly has never offered any mathematical, empirical, or

logical basis for assuming that thresholds derived from an analysis of Normal

distributions have any meaning for non-Normal distributions.  Commerce has

never addressed the fact that Normal distributions — whether based on a

sample or encompassing an entire population — have very different

mathematical characteristics than non-Normal distributions, and that those

different mathematical characteristics have a direct impact on the meaning of

the calculated $d$ statistic.  Appellees' briefs simply ignore this fundamental

flaw in Commerce's reasoning.

> 2.  *The Evidence Confirms that There Are No*
>     *Discernible Differences in the Prices for*
>     *Products that, According to Commerce, Show*
>     *a "Large" Effect Size under Its Cohen's d Test*

In our initial brief, we presented two diagrams of the price data for the

very first customer (number 102020) analyzed by Commerce's margin

calculation program, and the first product (control number 1-03-03-06-1) that

was found to pass the Cohen's $d$ test for that customer.  The first diagram

showed how the prices would appear if the data had satisfied the requirements

_____

[50] SeAH's September 9, 2015, Case Brief at 31, n.49 (Appx6372).

- 25 -

set forth in Professor Cohen's text.  In particular, we showed that two Normal

distributions with the means identified by Commerce ([number] for customer

102020 and [number] for the base group), with variances equal to the pooled

standard deviation of [number] calculated by Commerce, and with a large

enough number of data-points would look like this:

Normal Distributions with the
Same Means and Standard Deviations
as the "Target" and "Base" Groups for U.S. Sales
of Control Number 1-03-03-06-1 to Customer 102020
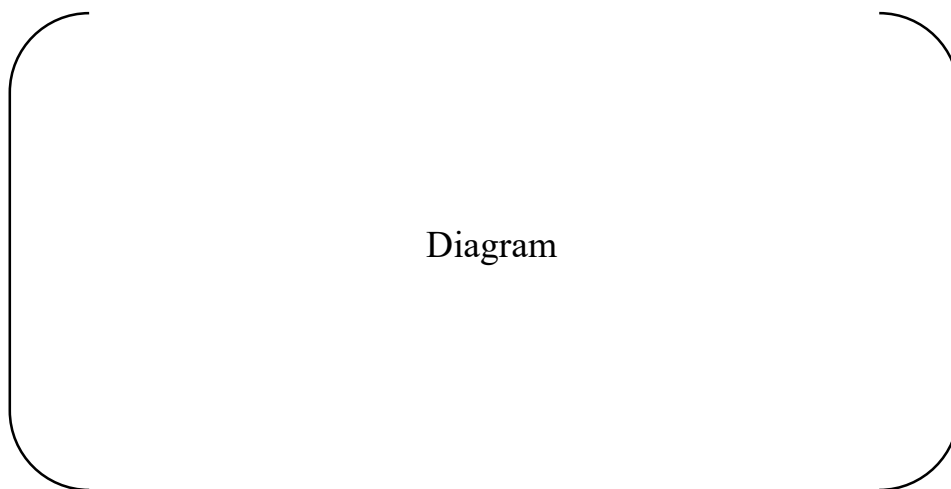
Diagram

As we noted, visual examination of the diagram would appear to show a

"pattern" of price differences between the two groups, with the prices for

customer 102020 appearing to be markedly *lower* than those of the base

group.

In our second diagram, we showed the actual dispersion of the individual

prices for sales to the target and base groups (by date of sale), as follows:

Actual Prices for U.S. Sales
in the "Target" and "Base" Groups for Sales
of Control Number 1-03-03-06-1 to Customer 102020

Diagram

As we noted, visual inspection of the actual individual sales prices in this

diagram does not reveal any actual pattern by customer.  The sales of control

number 1-03-03-06-1 to customer 102020 are clustered precisely in the

middle of the prices for sales of that product to other customers.[51]

---

[51] Numerical analysis of the actual prices reveals that there were 13 sales to the base group at prices below the lower of the two sales to customer 102020; 17 sales to the base group at prices above the higher of the two sales to customer 102020; and 6 sales to the base group at prices in-between the prices to customer 102020.

The Government objects to our first diagram on the grounds that it does not reflect SeAH's actual products for the product and customers in question.[52] The purpose of their objection is, however, unclear. We presented the first diagram as a *counterfactual* showing the situation that would exist if Professor Cohen's requirements had been satisfied (and for which a finding of a "large" effect size might have been appropriate). The fact that SeAH's actual prices did not follow that pattern was precisely the point.

The Government also objects to our second diagram on the grounds that it improperly assumes that a "large" difference in prices would be "visually discernible." According to the Government, "Many types of information cannot even be visually observed." The Government further contends that "for IQ scores of individuals or prices, one cannot visually see such data."[53] But those arguments are simply incoherent.[54]

---

[52] Government Brief at 49.

[53] *Id* at 49-50.

[54] Welspun also contends that the absence of any pattern of significant price differences depicted in our second diagram simply means that the variances in the prices was small. Welspun Brief at 49. According to Welspun, any distortion caused by a finding of a large Cohen's *d* from such small variances would be remedied by application of the "meaningful difference" test. *id.,* at 51-52. However, as demonstrated above, the "meaningful difference" test

*(footnote continued on following page)*

To begin with, Professor Cohen himself stated that his "medium" threshold was intended to identify differences that were "large enough to be *visible* to the naked eye," and his "large" threshold was intended to identify effects that were "grossly perceptible."[55]  Since a "medium" effect size is "visible to the naked eye," and a "large" effect size is even bigger than that, the Government's claim that "large" differences are not "visually discernible" simply cannot be reconciled with Professor Cohen's own description of his thresholds.

More generally, the Government's suggestion that information other than physical characteristics cannot be "visually observed" is refuted by an entire body of literature describing how such data can be visualized.[56]  Indeed, the academic literature considered by Commerce specifically referred to the

---

*(footnote continued from previous page)*
does not always remedy such issues.  And, more importantly, Welspun has misrepresented the problem identified by our second diagram:  The problem is not that the difference in prices is small but, instead, that there is no consistent difference in the prices for the specific customer and for the base group.

[55] *See* Cohen at 26-27 (Appx3770).

[56] *See., e.g.,* E. Tufte, THE VISUAL DISPLAY OF QUANTITATIVE INFORMATION (2d ed. 2001).

visual analysis of data.[57]  The Government's bizarre suggestion that data

concerning non-physical characteristics cannot be visualized demonstrates

only its desperation to dispute the irrefutable fact that there is no discernible

"pattern" or "significant difference" in the actual data for SeAH's sales

prices.

> 3.  *The Government's Depiction of*
> *Differences in the Average Prices for*
> *SeAH's Sales Is Deliberately Misleading*

Despite its assertion that prices cannot be visualized,[58] the Government

also contends that a significant difference in prices can be seen for the

customer and product referenced in our initial brief by a simple comparison of

the average prices for the two products.  In support of that claim, the

Government has submitted a bar chart that purports to compare the average

prices for the "test" customer and the "base group" for control number

---

[57] *See. e.g.,* Algina, Kesselman, Penfield, *An Alternative to Cohen's Standardized Mean Difference Effect Size: A Robust Parameter and Confidence Interval in the Two Independent Groups Case*, 10 PSYCHOLOGICAL METHODS 317, 319 and 327 (2005) (Appx4353, 4355, 4363); Li, *Effect Size Measures in a Two-Independent-Samples Case with Nonnormal and Nonhomogeneous Data*, 48 BEHAV RES 1560, 1561 (2018) (Appx4559-4560).

[58] Government Brief at 49-50.

1-03-03-06-1.  According to the Government, this bar chart shows a grossly

perceptible difference in the average prices of the two groups.[59]

It should be noted, however, that the Government's has truncated the

y-axis of its chart — starting its bars at a baseline of 600, rather than 0.  Such

truncation results in a gross exaggeration of the differences between the two

bars.  Indeed, a bar chart with a baseline different from zero has been

described by one expert on data visualization as "the most conspicuous trick

to distort your perception of numbers."[60]  Other experts have decried the

truncation of bar charts in such a manner as "biased," "dishonest,"

"deceptive," "lying with statistics," and "the worst of crimes in data

visualization."[61]  The Government's reliance on a chart employing such

trickery does not establish that a perceptible difference actually existed.[62]

---

[59] Government Brief at 51.

[60] *See* Cairo, HOW CHARTS LIE (2019), at 13.

[61] *See* Correll, Bertini, and Franconeri, *Truncating the Y-Axis: Threat or Menace?*, PROCEEDINGS OF THE 2020 CHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, available at « https://arxiv.org/pdf/1907.02035.pdf ».

[62] The Government contends that the use of a truncated Y-Axis is consistent with SeAH's presentation of the scatter-plot of actual prices for the relevant control number.  However, the issue of truncating the Y-Axis does not arise in graphs, such as scatter-plots, that show trends in data and do not attempt to draw conclusions about differences in absolute magnitude.  *See* Correll *et al.* at 3. ("{E}xisting guidelines for the design of line graphs and scatterplots

In any event, it is ironic that the Government has suggested that a simple comparison of average prices is sufficient to identify a significant difference in prices. After all, the entire purpose of the DPA is to determine whether comparisons are distorted by the use of averages. Comparing averages that mask the trends in individual prices to determine whether averages mask the trends in individual prices is inherently illogical and unreasonable.

E.    *Commerce's Statistics Regarding the Impact of the DPA on Its Results Are Misleading*

As noted in our initial brief, Commerce asserted, and the CIT appears to have agreed, that the reasonableness of Commerce's methodology could be assessed based on statistics concerning the percentage of cases in which the DPA affected the result. In response, we noted (1) that the "reasonableness" of the number of cases affected cannot be assessed in a vacuum, without knowing what the "correct" percentage should have been; (2) that Commerce's statistics are irrelevant to the situation presented in this case, because they included the results for investigations in which Commerce applied a different standard for measuring whether a "meaningful difference" existed; and (3) that an analysis of investigations in which Commerce applied

---

*(footnote continued from previous page)*
focus on making the overall trend as visible (and decodable with the least error) as possible.").

the same "meaningful difference" test that was applied in this investigation would have found that Commerce's DPA increased the dumping margins from *de minimis* to above *de minimis* in more than 50 percent of such investigations.[63]

Appellees contend that Commerce's analysis of the impact of the DPA on its cases was reasonable, because it included all investigations during the period considered.[64] But that argument ignores the key point: Commerce does not have a single DPA that it applies in all cases. Instead, it has two different DPAs — one that applies when the A-to-A margin would be *de minimis* (in which case, any change from below *de minimis to* above *de minimis* is considered to constitute a "meaningful difference"), and a second that applies when the A-to-A margin would be above *de minimis* (in which case, an increase in the dumping margin of 25 percent is required to find a "meaningful difference").

In this case, SeAH's dumping margin under the A-to-A methodology would have been 1.97 percent. Commerce therefore applied its first version of the DPA. Under that version, in order for a "meaningful difference" to be

---

[63] *See* SeAH's Initial Brief at 38-43.

[64] *See* Government Brief at 56-58, Welspun Brief at 59-61.

found, all that was required was that the dumping margin increase from 1.97 to 2.00 percent. In other words, an increase of 0.03 percentage points (or 1.5 percent of SeAH's A-to-A margin) would have been sufficient to cause Commerce to apply the A-to-T methodology. The effect of the DPA on other cases — where the dumping margins would have been above *de minimis* under any methodology, and where Commerce required an increase in the dumping margin of 25 percent — says nothing about the effect of the DPA in the form in which it was applied to SeAH.

In the end, however, the entire discussion is irrelevant, because the reasonableness of a methodology cannot be assessed based on its impact on the results in the absence of an objective baseline. The chance of rolling a "lucky seven" with two six-sided dice is one in six (or 16.67 percent). But we cannot imagine that even the most deferential Court would allow Commerce to decide whether to apply the A-to-A or A-to-T methodology just by rolling the dice and applying the A-to-T methodology whenever it rolled a seven. The fact that 16.67 percent may fall within some judge's preconceived notion of an "acceptable" range does not make rolling the dice a reasonable methodology.

As we have explained at length, Commerce does not apply its Cohen's *d* test in a statistically-sound manner. As a result, the outcome of its analysis is

essentially as random as a dice roll.  Figures about the percentage of cases

affected by Commerce's analysis cannot transform its random test into a

reasonable implementation of the statutory requirements.

<div align="center">CONCLUSION</div>

In light of the foregoing, we respectfully request that the Court reverse

the decision by the CIT and remand this case with instructions requiring

Commerce to recalculate the dumping margins for SeAH using an average-to-

average comparison.

Respectfully submitted,

/s/Jeffrey M. Winton

Jeffrey M. Winton
Jooyoun Jeong

WINTON & CHAPMAN PLLC
1100 13th Street, N.W., Suite 825
Washington, D.C.  20005
(202) 774-5500

Attorneys for SeAH Steel Corporation

January 8, 2024

FORM 19. Certificate of Compliance with Type-Volume Limitations

Form 19
July 2020

# UNITED STATES COURT OF APPEALS
# FOR THE FEDERAL CIRCUIT

## CERTIFICATE OF COMPLIANCE WITH TYPE-VOLUME LIMITATIONS

**Case Number:** 2023-1663

**Short Case Caption:** Stupp Corporation v. U.S.

---

**Instructions:** When computing a word, line, or page count, you may exclude any items listed as exempted under Fed. R. App. P. 5(c), Fed. R. App. P. 21(d), Fed. R. App. P. 27(d)(2), Fed. R. App. P. 32(f), or Fed. Cir. R. 32(b)(2).

---

The foregoing filing complies with the relevant type-volume limitation of the Federal Rules of Appellate Procedure and Federal Circuit Rules because it meets one of the following:

☑ the filing has been prepared using a proportionally-spaced typeface and includes __6,977__ words.

☐ the filing has been prepared using a monospaced typeface and includes _____ lines of text.

☐ the filing contains _____ pages / _____ words / _____ lines of text, which does not exceed the maximum authorized by this court's order (ECF No. _____).

Date: __01/08/2024__

Signature: /s/ Jeffrey M. Winton

Name: Jeffrey M. Winton

**FORM 30. Certificate of Service**

<div align="right">

**Form 30**
**July 2020**

</div>

# UNITED STATES COURT OF APPEALS
# FOR THE FEDERAL CIRCUIT

## CERTIFICATE OF SERVICE

**Case Number**  2023-1663

**Short Case Caption**  Stupp Corporation v. U.S.

---

**NOTE:** Proof of service is only required when the rules specify that service must be accomplished outside the court's electronic filing system.  See Fed. R. App. P. 25(d); Fed. Cir. R. 25(e).  Attach additional pages as needed.

---

I certify that I served a copy of the foregoing filing on  01/08/2024

by  ☐  U.S. Mail  ☐  Hand Delivery  ☐ Email  ☐ Facsimile
☑  Other:  Password-protected link sent via Email

on the below individuals at the following locations.

| Person Served | Service Location (Address, Facsimile, Email) |
|---|---|
| Robert R. Kiepura | robert.kiepura@usdoj.gov |
| Jeffrey David Gerrish | jgerrish@schagrinassociates.com |
|  |  |
|  |  |
|  |  |

☐  Additional pages attached.

Date: 01/08/2024

Signature:  /s/ Jeffrey M. Winton

Name:  Jeffrey M. Winton

# UNITED STATES COURT OF APPEALS
# FOR THE FEDERAL CIRCUIT

## CERTIFICATE OF CONFIDENTIAL MATERIAL

**Case Number:**  2023-1663

**Short Case Caption:**  Stupp Corporation v. U.S.

---

**Instructions:** When computing a confidential word count, Fed. Cir. R. 25.1(d)(1)(C) applies the following exclusions:

- Only count each unique word or number once (repeated uses of the same word do not count more than once).

- For a responsive filing, do not count words marked confidential for the first time in the preceding filing.

The limitations of Fed. Cir. R. 25.1(d)(1) do not apply to appendices; attachments; exhibits; and addenda. *See* Fed. Cir. R. 25.1(d)(1)(D).

---

The foregoing document contains $\underline{5}$ number of unique words (including numbers) marked confidential.

☐ This number does not exceed the maximum of 15 words permitted by Fed. Cir. R. 25.1(d)(1)(A).

☑ This number does not exceed the maximum of 50 words permitted by Fed. Cir. R. 25.1(d)(1)(B) for cases under 19 U.S.C. § 1516a or 28 U.S.C. § 1491(b).

☐ This number exceeds the maximum permitted by Federal Circuit Rule 25.1(d)(1), and the filing is accompanied by a motion to waive the confidentiality requirements.

Date: 01/08/2024         Signature: /s/ Jeffrey M. Winton

                         Name:       Jeffrey M. Winton